128–133 *Nucleic Acids Research, 1999, Vol. 27, No. 1*

MitBASE: a comprehensive and integrated mitochondrial DNA database

M. Attimonelli[1,*], N. Altamura[2], R. Benne[3], C. Boyen[4], A. Brennicke[5], A. Carone[1], J. M. Cooper[6], D. D'Elia[7], A. de Montalvo[8], B. de Pinto[2], M. De Robertis[1], P. Golik[9], J. M. Grienenberger[10], V. Knoop[5], C. Lanave[2], J. Lazowska[9], A. Lemagnen[9], B. S. Malladi[11], F. Memeo[1], M. Monnerot[9], S. Pilbout[9], A. H. V. Schapira[6,12], P. Sloof[3], P. Slonimski[9], K. Stevens[6,12] and C. Saccone[1]

[1]Dipartimento di Biochimica e Biologia Molecolare, Università degli Studi di Bari, 70126 Bari, Italy, [2]Centro di Studio sui Mitocondri e Metabolismo Energetico, CNR, 70126 Bari, Italy, [3]Vakgroep Biochemie, AMC Meibergdreef, Amsterdam, The Netherlands, [4]CEOBM, Station Biologique CNRS, Roscoff, France, [5]Ulm University, D-89069 Ulm, Germany, [6]University Department of Clinical Neurosciences, Royal Free and University College Medical School, UCL, London, UK, [7]Area di Ricerca, CNR, 70126 Bari, Italy, [8]Departamento de Biología Molecular, Hospital Universitario 'Marqués de Valdecilla', Universidad de Cantabria, 39011 Santander, Spain, [9]Centre de Genetique Moleculaire, CNRS, 91198 Gif-sur-Yvette, France, [10]IBMP, CNRS, 67084 Strasbourg, France, [11]EBI Hinxton Hall, Hinxton, Cambridge, UK and [12]University Department of Clinical Neurology, Institute of Neurology, London, UK

Received September 2, 1998; Accepted September 7, 1998

ABSTRACT

MitBASE is an integrated and comprehensive database of mitochondrial DNA data which collects all available information from different organisms and from intra-specie variants and mutants. Research institutions from different countries are involved, each in charge of developing, collecting and annotating data for the organisms they are specialised in. The design of the actual structure of the database and its implementation in a user-friendly format are the care of the European Bioinformatics Institute. The database can be accessed on the Web at the following address: http://www.ebi.ac.uk/htbin/Mitbase/mitbase.pl . The impact of this project is intended for both basic and applied research. The study of mitochondrial genetic diseases and mitochondrial DNA intraspecie diversity are key topics in several biotechnological fields. The database has been funded within the EU Biotechnology programme.

INTRODUCTION

Mitochondrial DNA (mtDNA) is an essential component of all eukaryotic cells. It ensures consistency of function (cellular respiration and oxidative phosphorylation) despite the great diversity of genome organisation. Depending on the organism, the amount of information contained in the mtDNA varies. Still, the mtDNA of virtually all organisms contains genes encoding ribosomal RNAs, tRNAs and proteins which are part of the enzyme complexes of the inner mitochondrial membrane and participate in oxidative phosphorylation (1). In some taxonomic groups, most notably in plants and protists, mitochondrial genes with other functions are found, e.g. encoding ribosomal proteins, components of cytochrome c biogenesis or proteins with an as yet unidentified function (2). Moreover nuclear DNA codes for several components interacting with the mitochondrial DNA. Thus, studies on mitochondrial DNA are important to unravel the interaction of the two genomes, one of the most intriguing and important aspects of the eukaryotic cell.

The advent of sequencing technologies and their recent improvement have resulted in the production of mtDNA sequences for a large number of species and variants. Since the early 1980s, the determination of several complete animal mitochondrial genomic sequences [human (3), mouse (4), cow (5) and rat (6) genomes] has produced a great quantity of data and hence a lot of information is available, but, as in a puzzle, the dispersed pieces need to be assembled. The vast differences in mitochondrial genome organisation and mode of gene expression observed between the taxonomic groups has made this a very difficult task. Ever more problematic is the storage of all the data in such a way that they can be retrieved and analysed.

To alleviate some of these problems, a project was initiated as a collaborative effort of several European groups aiming at the construction of a comprehensive and integrated database, collecting information on mtDNA from all organisms which contain mitochondria in their cells: MitBASE.

The most important goal of the MitBASE project is to provide the scientific community with all available information both on different organisms and on intraspecie variants and mutants

*To whom correspondence should be addressed. Tel: +39 080 548 2130; Fax: +39 080 548 4467; Email: marcella@area.ba.cnr.it

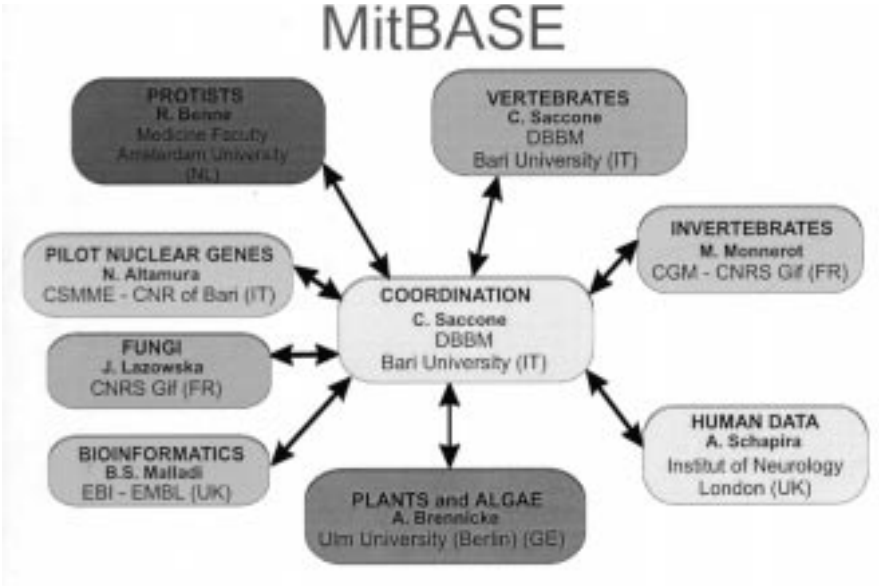

Figure 1. MitBASE network. The figure reports MitBASE network organisation. For each node, the name and the affiliation of the official scientist responsible within the EU project is reported.

collected by experts thus assuring high quality standards and accuracy in the annotation, validation and structuring of the data.

The present paper describes the current state of the MitBASE project.

MitBASE NETWORK

MitBASE was started as a project funded within the EU Biotechnology programme co-ordinated by C. Saccone.

This project is a collaboration among researchers expert in human, vertebrate, invertebrate, protist, fungal and plant mtDNAs. A Pilot node aimed at collecting nuclear genes involved in mitochondrial biogenesis in *Saccharomyces cerevisiae* is also a part of the project. The informatic part of the project is managed at the EBI, the European Bioinformatics Institute. The location of the eight nodes is reported in Figure 1.

MitBASE is defined as an integrated database because it is a collection of specialised databases, one per node, differentiated by the specificity of the data each node has to manage, the guiding thread being the mtDNA.

One of the special features of MitBASE is the presence in it of '*variants*'. A *variant* can be defined as any fragment where nucleotide differences have been detected as compared with a reference sequence which can be associated to a real individual or a synthetic fragment resulting from the consensus between a set of sequences. At present *variants* are available in the Human database (associated with mitochondrial pathologies and human diversity studies), in the Vertebrate database (predominantly in the class Fish as a result of stocking analyses), in the Invertebrate database and in the Fungal database (mutants having a clear genotype/phenotype relationship). The presence of *variants* in MitBASE is one of its most important features which diversifies MitBASE from GOBASE (7), the Organelle Genomes Database. Another salient feature of MitBASE is the storage in it of data related with the *editing process* occurring in the plant and protist species.

DATA STRUCTURES

A specific data-field set has been defined, one per node, describing the classes and the objects to be stored in MitBASE, based on which the data structures for each node have been defined.

Data are locally managed in different ways. Some nodes have collected data supported by commercial Database Management Systems (Microsoft Access for the Human and Vertebrates Nodes, and for the Fungal Mutant data; FileMaker Pro for the Pilot Node). Other nodes (Invertebrate, Fungal for the non-mutant section, Plant and Protist) have simply organised their data in a flatfile format derived from the primary databases.

The data organised in these structures are periodically sent to the EBI where they are stored in a centralised ORACLE based database.

Within ORACLE the data are stored in diversified structures reproducing the local organisation. The integrated MitBASE database can be queried at the EBI WWW site through different approaches or can be released in a flatfile format suitable for management with any other biological database query system [i.e. SRS (8)].

DATA SOURCES

Data sources in MitBASE are the primary databases [EMBL Data Library (9) and GenBank (10)], literature through bibliographic databases (Medline, Current Contents and Current Advances in Genetics and Molecular Biology) and personal communications.

Data are locally collected at each node using different approaches. The common policy adopted at each node in data collection is an accurate revision of the data, which guarantees the quality of the data stored in MitBASE.

Revised data are enriched with value added information derived from other sources or from specific analyses performed by the experts working on the project.

MitBASE NODES DESCRIPTION

A concise description of the peculiarity associated with each node is reported below.

The Human database

The need to create a separate node for human data is justified by the great quantity of human mtDNA data related to studies on population genetics and to studies about the relationship between human mtDNA alterations and diseases.

Each entry in the Human dataset is a *variant* related to the complete human mitochondrial genome sequence reported by Anderson *et al*. in 1981 (3). Each variant can be associated to more individuals, identified by family codes whose mtDNA has been extracted from different tissues. To each *variant*, information is added to better describe the individual: geographical and linguistic classification, clinical, biochemical and histopathological features (11).

Two datasets have been defined for the Human node: the Molecular dataset and the Clinical dataset. Data entry is thus performed by subnodes, one based at the Institute of Neurology in London (UK) and the other at the Department of Biochemistry and Molecular Biology in Bari (Italy). The Bari group is in charge of the management of the human mtDNA molecular information while the London group manages the clinical data.

Human data can be queried through the MitBASE simple query system (see below) and by using SRS in the section 'Mutation Databases' (http://srs.ebi.ac.uk/srs5/).

The Vertebrate database

This node collects prevalently data from primary databases (EMBL and GenBank). The major aim of this node is to revise the data already in the primary databases, to eliminate redundancies and to add information not present in the primary databases such as the Geographic origin of the samples, the methods used in the production of the sequences, the methods used for the theoretical analyses of the sequences and, for genes coding for proteins and tRNAs, the links to a database collecting the multiple alignment of the metazoan mitochondrial genes coding for proteins and tRNAs: AMmtDB (12). In order to save time and ensure safe data submission the data already in the primary databases are checked and revised through a WWW interface developed at the EBI and available through the MitBASE site. The revised data are automatically stored in the centralised database.

The new data are stored locally in the Microsoft Access database and hence are exported through tables in the ORACLE centralised database.

In order to store *variants* data in this database the software VARCLU (B.S.Malladi, G.Grillo and A.Carone, manuscript in preparation) has been developed to support the annotators in the selection of the reference sequence and in the detection of variation events (point substitutions, insertions and deletions). Data input for this software is a CLUSTAL (13) multiple alignment.

The Invertebrate database

The Invertebrate database shares its aims with the Vertebrate database, i.e. checking the accuracy of information already present in the EMBL and GenBank databases and add information (mainly: geographic origin, details on organisms, methods used for establishing sequences). When this information is not available from literature, the authors are directly asked to provide it. Data relevant to polymorphisms are also enclosed as variants.

Due to the great diversity in gene organisation in invertebrates, special attention is given to gene recognition and location; details are added for tRNA and rRNA gene description.

Work developed for establishing Mitotool, a tool for queries on sequences and dynamic alignments (A.Lemagnen, manuscript in preparation), is proving helpful in some informatic development of MitBASE.

The Fungal database

The Fungal node collects fungal wild-type mitochondrial sequences including those of ascomycetes and filamentous ascomycetes and fungal mitochondrial mutations/variants. The mitochondrial genomes of fungi, in contrast to the highly compact genomes of multicellular organisms, are extremely heterogeneous in size (from 9 to 101 kb) and shape (circular and linear). For these reasons the overall genetic organisation of fungal mtDNA cannot be related to the structure of a reference genome as happens, for instance, in mammalian mitochondrial genomes. In addition, many fungal intraspecie strains in *S.cerevisiae* exhibit polymorphism ranging from small differences in coding or non-coding regions to large variations in relevant genetic elements, mainly introns, which can be present or absent. Finally, the presence of open reading frames in the intronic regions has induced fungal annotators to assign standard names to fungal intronic open reading frames.

Only a small part of the fungal sequence data is available in primary databases; the majority of them and namely mutants are in original papers, in theses or are unpublished, which causes major difficulties in data capturing.

The Plant and Alga database

Particular to land plant mtDNA are the differences between genomic and mRNA sequences introduced by RNA editing (pyrimidine conversions), the frequent genomic recombinations leading to alternative or coexisting gene arrangements, the presence of promiscuous sequences originating in the nuclear and chloroplast DNA and the fragmentation of some genes requiring trans-splicing for mRNA maturation. Information on most of these sequence features was hitherto only available through the original literature references. While annotated in the primary databases, they were not accessible to computer queries and frequently contained errors due to a lack of standardisation in annotation. In the process of data entry revision, the errors in annotation have been corrected (and feedback was also given to the primary databases and authors) and the new features have been introduced into MitBASE in a standardised way to allow querying in a sequence oriented way.

The Protist database

The protist group is not phylogenetically equal to that of plants, animals and fungi because of the enormous cytological, organisational and molecular diversity of these organisms. This diversity is also reflected in protist mtDNAs, which can be either linear or circular, sometimes, as in kinetoplastids, even consisting of a network of catenated circles, and which are extremely heterogeneous in gene content and size. An additional complication is the occurrence of different forms of RNA editing. Gene-encoded

sequences can be altered by insertion/deletion of Us (in kinetoplastids), insertion of C, U and various dinucleotides, substitution of C by U (in Physarum), and by various other nucleotide substitutions (in Acanthamoeba). In kinetoplastids, the information for the edited RNA sequences is provided by small guide (g)RNAs, which are encoded (mostly) in the minicircle component of the mtDNA. Therefore, the Protist node not only collects, revises and organises all protist mtDNA sequences, but also edited RNA and gRNA sequences. Whenever such information is available, alignments of edited RNA:DNA and edited RNA:gRNA are provided based on the U insertion/deletion database (http://www.lifesci.ucla.edu/RNA/trypanosome/database.html ; L. Simpson, UCLA), which contains such alignments for four and two species, respectively. If known, these alignments are directly linked to the genomic location of the guide RNAs and the sequence of the corresponding minicircles, using information from the primary databases, but also from the Guide RNA database (A. E. Souza, S. Hinz and H. Ulrich, Munchen http://www.biochem.mpg.de/~goeringe/gRNA/gRNAseqs.html) and the minicircle database (D. C. Barker, S. Brewster and M. Aslett, Cambridge, UK; http://www.ebi.ac.uk/parasites/kDNA/Source.html).

Last but not least, edited RNA:DNA alignments for the other organisms with mitochondrial RNA editing are provided.

The Pilot database

The Pilot database contains a compilation of nuclear genes accurately selected on the basis of their direct or indirect involvement in the biogenesis of functional mitochondria and its regulation. The information data set has been defined such as to integrate and complement basic molecular and genetic data from primary nucleic acids databases (EMBL and GenBank) and from yeast databases [mainly MIPS (14) and YPD (15)] with information related to peculiar aspects of the mitochondrial research, particularly the mitochondrial phenotype of the gene knockout, as a result of an extensive screening in literature. Genes have been classified according to the mitochondrial process in which their products participate. An elaborated and specific query system has been developed for this database and it can be accessed through the MitBASE home page (de Pinto *et al.*, submitted for publication). The aim of this database is to provide the scientific community with a basic plan of the nuclear contribution to the mitochondrial biogenesis such as to constitute a reference model with which to compare other organisms including man.

THE MitBASE HOME PAGE

The integrated database and any other information related to it, and more generally to mtDNA, is available through the MitBASE home page. After a brief description of the project and its goals, the list of the functions implemented through the home page is illustrated. Each element of the list works as a button that, if clicked on, starts functions or allows to view specific documents. Some of the available functions are described here below.

Submission data interface

This interface is adaptable to submissions by different nodes. It facilitates retrieval of information related to entries already available in primary databanks (e.g., EMBL) and provides their output in an editable interface so that, after correction, data can be sent to EBI and included in MitBASE. Such an interface has been developed to avoid the 'time-consuming' work of submission of information mostly common to all the nodes in MitBASE which consists of citations, taxonomy, cross-references and sequence data.

Gene names classification

One of the problems in MitBASE is related to the naming of mitochondrial genes. In primary databases the same gene is named in many different ways both because different taxonomic groups have historically assigned to the same gene a different name and because of typographic differences in spelling the gene name by different annotators. This would have implied that searching for a given gene in any taxonomic group would have caused false results because different entries from the same genes in MitBASE had been classified with different gene names. Therefore, a standard classification has been adopted and, in order to maintain the memory of the gene name assigned by the original author, the alt_gene field has been appended to the gene field both in the structure and in the output flatfile. The standard gene name classification is based on GOBASE (7) gene names with some modifications as implemented in KEYnet (the KEYnet Project is one of the activities of the EU project 'Provision of the EMBL Data library' coordinated at EBI by G. Cameron and funded under the BIOTECH program. Keynet is developed at the CNR Research Area in Bari and is the responsibility of M. Attimonelli; submitted to NAR special issue, http://www.ba.cnr.it/keynet.html). This classification is available in document format from MitBASE homepage and can be used as a guide for gene name usage. MitBASE gene names classification will soon include fungal intronic open reading frame names as they have been assigned by the Fungal node team. The Fungal node has defined a clear and unified system for naming intronic open reading frames and the final document describing such a nomenclature will be soon available on MitBASE home page.

MitBASE query systems

Data in MitBASE can be queried using different approaches. In particular a '*simplified query system*' allowing to search data from all the nodes has been implemented at the EBI site. Moreover '*elaborated query systems*' specific for each node are under development. The simplified query system allows to search data according to gene and species grouped by taxa, corresponding to the different MitBASE nodes. Searching by species has been implemented at the EBI in collaboration with the Invertebrates node and will allow to browse the NCBI Taxonomy Classification for organisms with mitochondria. The selected data can be viewed in tables based on the ORACLE structure and for some taxa (human and vertebrates) also in the MitBASE flatfile format. For each node in MitBASE an elaborated query system is being developed. In principle it should allow to search data by combining, through the logical operators, all the criteria corresponding to the fields defining the node-specific structure. Selected data will be released in MitBASE flatfile format and managed in tables in order to produce graphs and statistical analysis. At present, the Pilot elaborated query system is fully implemented (de Pinto *et al.*, submitted to NAR) and can be accessed through the MitBASE home page or directly at the following address: http://www3.ebi.ac.uk/Research/Mitbase/mitbiog.pl . The Plant node elaborated query system has been designed and is currently being tested. It is available to the following address: http://tonic.ebi.ac.uk:8889/mitbase/plsql/pla_qry.pla_show_qry_opts . The MitBASE plant

Query System (S.Malladi and V.Knoop, manuscript in preparation) allows queries like: 'List all C to U RNA editing sites in atp6 genes of plants that are preceded by the sequence motif UC and followed by GGA or which trans-splicing group II introns are known from nad2 genes of dicotyledons' or 'are any stretches of promiscuous chloroplast DNA larger than 560 bp known in Petunia, maize and tobacco?'. The elaborated query systems for the other nodes are being designed.

MitBASE entry's status

This function allows to have an overview of the number of entries stored in MitBASE for each node (Table 1). The total number of entries available in MitBASE at August 1998 was 10 094.

Table 1. MitBASE content at August 1998

Node	Entries in Database
Human	4756
Vertebrate	2014
Plant and Alga	381
Invertebrate	1631
Protist	320
Fungal	498
Pilot Nuclear Genes	494

Link to mitochondrial translation tables

In order to allow MitBASE annotators and users to check the organism specificity of the mitochondrial genetic codes, a link to the translation tables available at the EBI is being set up.

Pointers to mitochondrion related information WWW sites

This function allows browsing other available mitochondrial databases [GOBASE (7), MmtDB (12), MITOMAP (16), MITODAT (http://www-lmmb.ncifcrf.gov/mitoDat), compilation of human mtDNA control region sequences (17)], and access to other WWW sites where relevant information on the mitochondrion can be obtained.

THE MitBASE FLATFILE FORMAT

The database is presently available on the WWW site at the EBI. It will be released also in flatfile (ff) format in order to allow its usage with other systems available worldwide. The users of the present database are kindly invited to cite this article. The flatfile format has been defined by a task force. Although the peculiarities of the dataset at each node require specific rules for the structuring of their ff, some general criteria have been identified and the general scheme is reported in Figure 2. Each entry in the MitBASE ff is associated to a nucleotide sequence coding for a complete mtDNA genome or for part of it. Moreover, in the Human, Vertebrate, Invertebrate and Fungal databases, entries associated to variants or mutants are also generated. Finally, for the Pilot database each entry refers to a nuclear gene involved in the mitochondrial biogenesis.

```
ID    MitBASE Identifier; Molecule type; Sequence length;
DE    Free text description
OS    Organism name
OC    Taxonomic classification
RX    Medline Crossreferencing
RN    Reference numbering
RA    Authors list
RL    Full reference
DR    Database name; Accession number; Entry name
FH    Key           Location/qualifiers
FT    CDS           x..y
FT                  /gene="gene name"
FT                  /alt_gene="synonymous gene name"
FT                  /transl_code="#"
FT                  /product="protein name"
FT                  /note="adding value information"
FT    intron        x..y
FT                  /gene="gene name"
FT                  /number="#"
FT    exon          x..y
FT                  /gene="gene name"
FT                  /number="#"
FT    seq_blocks    x..y
FT                  /note="CSB"
      or
FT                  /note="TAS"
      or others
FT    ..........
SQ
1     acatttg.......
//
```

Figure 2. MitBASE general Flatfile format. The identification (ID) line contains the unique identifier coded according to the following rules MTXX##### i.e. MT for all nodes, XX a two letter code identifying the node (VR = vertebr, IN = invert, PD = plants dicots, PM = plants monocots, PA = plants algae, PR = protists, HS = human, FU = fungi, PN = pilot) and a five digit code automatically generated during data storage. The description (DE) lines contain free text describing species name, sequence function and other peculiarities. The organism classification (OC) lines are automatically generated by NCBI Taxonomic classification. The reference (RX, RN, RA, RL) and the cross-referencing (DR) lines are coded according to EMBL format. Finally, for the features table (FT) lines, EMBL format has been adopted as a model although new qualifiers have been introduced and new rules have been fixed. As far as the translation qualifier for CDS features lines is concerned, sequence translation has not been considered and only transl_code qualifiers have been included. The transl_table refers to the translation tables available on the MitBASE home page. MitBASE specific qualifiers and two letter codes have been adopted depending on the taxonomic group (not shown).

ACKNOWLEDGEMENTS

We are grateful to Dr M. di Tommaso and to the EBI database staff for the informatic support to the design and implementation of the database. This work has been funded under the EU Biotechnology programme, contract number: BIO4 CT950160.

REFERENCES

1 Saccone,C. (1994) *Curr. Opin. Genet. Dev.*, **4**, 875–881.
2 Gray,M.W., Lang,B.F., Cedergren,R., Golding,G.B., Lemieux,C., Sankoff,D., Turmel,M., Brossard,N., Delage,E., Littlejohn,T.G., Plante,I., Rioux,P., Saint-Louis,D., Zhu,Y. and Burger,G. (1998) *Nucleic Acids Res.*, **26**, 865–878.
3 Anderson,S., Bankier,A.T., Barrell,B.G., de Bruijn,M.H., Coulson,A.R., Drouin,J., Eperon,I.C., Nierlich,D.P., Roe,B.A., Sanger,F., Schreier,P.H., Smith,A.J., Staden,R. and Young,I.G. (1981) *Nature*, **290**, 457–465.

4 Bibb,M.J., van Etten,R.A., Wright,C.T., Walberg,M.W. and Clayton D.A. (1981) *Cell*, **26**, 167–180.
5 Anderson,S., Bruijn,M.H., Coulson,A.R., Eperon,I.C., Sanger,F. and Young,I.G. (1982) *J. Mol. Biol*., **156**, 683–717.
6 Gadaleta,G., Pepe,G., De Candia,G., Quagliariello,C., Sbisa,E. and Saccone,C. (1989) *J. Mol. Evol*., **28**, 497–516.
7 Korab-Laskowska,M., Rioux,P., Brossard,N., Littlejohn,T.G., Gray,M.W., Lang,B.F. and Burger,G. (1998) *Nucleic Acids Res*., **26**, 138–144.
8 Etzold,T., Ulyanov,A. and Argos,P. (1996) *Methods Enzymol*., **266**, 114–128.
9 Stoesser,G., Moseley,M.A., Sleep,J., McGowran,M., Garcia-Pastor,M., Sterk,P. (1998) *Nucleic Acids Res*., **26**, 8–15.
10 Benson,D.A., Boguski,M.S., Lipman,D.J., Ostell,J. and Ouellette,B.F. (1998) *Nucleic Acids Res*., **26**, 1–7.
11 Attimonelli,M., Calo,D., Cooper,J.M., de Montalvo,A., Licciulli,F., Sasanelli,D., Stevens,K., Malladi,B.S., Saccone,C. and Shapira,A.H. (1998) *Nucleic Acids Res*., **26**, 116–119.
12 Attimonelli,M., Calò,D., De Montalvo,A., Lanave,C., Sasanelli,D., Tommaseo Ponzetta,M. and Saccone,C. (1998) *Nucleic Acids Res*., **26**, 120–125.
13 Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) *Nucleic Acids Res*., **22**, 4673–4680.
14 Mewes,H.W., Hani,J., Pfeiffer,F. and Frishman,D. (1998) *Nucleic Acids Res*., **26**, 33–37.
15 Hodges,P.E., Payne,W.E. and Garrels,J.I. (1998) *Nucleic Acids Res*., **26**, 68–72.
16 Kogelnik,A.M., Lott,M.T., Brown,M.D., Navathe,S.B. and Wallace,D.C. (1998) *Nucleic Acids Res*., **26**, 112–115.
17 Handt,O., Meyer,S. and von Haeseler,A. (1998) *Nucleic Acids Res*., **26**, 126–129.